

DOCUMENT RESUME

ED 203 661

FL 012 285

AUTHOR DeMauro, Gerald E., Ed.
 TITLE The Impact of Bilingual Education on English Acquisition in New Jersey.
 INSTITUTION New Jersey State Dept. of Education, Trenton. Div. of Research, Planning, and Evaluation.
 PUB DATE Jan 81
 NOTE 63p.
 EDRS PRICE MF01/PC03 Plus Postage.
 DESCRIPTORS *Bilingual Education; Educational Assessment; Elementary Secondary Education; *English (Second Language); *Program Effectiveness; Second Language Learning; Spanish Speaking
 IDENTIFIERS Bilingual Programs; Language Assessment Battery; New Jersey

ABSTRACT

The Language Assessment Battery was administered as a pre- and post-test to 731 students (mostly Spanish-speaking) in bilingual programs in New Jersey. In all grades tested, results consistently showed that third-year bilingual students outscored first and second year students in English language skills. This supports the contention that bilingual education promotes English skills development. It was also demonstrated that bilingual students had made significant gains relative to English-speaking children in eight of the twelve grades tested. For example, the average post-test score of second grade students who had started bilingual programs in kindergarten surpassed the average English-speaking student's theoretical norm (54.3 as compared to 50.0). This finding argues in favor of early bilingual education intervention. Finally, it was discovered that, in general, elementary school students made much larger gains than high school students. Recommendations for adjustments to bilingual policy based on these findings are set forth. (Author/JB)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

THE IMPACT OF BILINGUAL EDUCATION
ON ENGLISH ACQUISITION IN NEW JERSEY

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

FRED G. BURKE
Commissioner of Education

GUSTAV H. RUH
Deputy Commissioner of Education

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Gerald E. DeMauro

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

PREPARED UNDER THE DIRECTION OF:
DIVISION OF RESEARCH, PLANNING AND EVALUATION

RITA J. CARNEY
Assistant Commissioner
Division of Research, Planning and Evaluation

DEVELOPED AND EDITED BY:

GERALD E. DEMAURO
Bilingual Evaluation Consultant
Division of Research, Planning and Evaluation

New Jersey State Department of Education
225 West State Street
Trenton, New Jersey 08625

- January 1981 -

PTM NO.

ED203661

FL012 285

STATE BOARD OF EDUCATION

P. Paul Ricci, President.....Vineland
S. David Brandt, Vice President.....Cherry Hill
Jack Bagan.....Salem
Anne S. Dillman.....Perth Amboy
Ruth H. Mancuso.....Glassboro
Katherine K. Neuberger.....Lincroft
(Representing State Board of Higher Education)
Sonia B. Ruby.....Atlantic City
Susan N. Wilson.....Princeton
Robert J. Wolfenbarger.....Summit

Fred G. Burke, Commissioner of Education
Secretary

T. Edward Hollander, Chancellor of Higher Education
Ex Officio

Distribution List

County Superintendents	21
District Superintendents	611
Department Heads	25
Division Heads	As requested
General public	As requested
Other State Offices (Governor's office, etc.)	As requested

TABLE OF CONTENTS

Executive Summary.....	1
Background.....	1
Purpose of Study.....	1
Procedures.....	11
Results.....	11
Conclusions and Recommendations.....	111
I. INTRODUCTION.....	1
A. History.....	1
B. Study Question: Can Children Learn English While Being Instructed in Their Native Language?.....	1
C. Research Designs.....	1
D. Study Limitation.....	2
II. METHODS AND PROCEDURES.....	3
A. Choice of Instruments.....	3
B. Determination of Sample Size.....	3
C. District Sampling.....	4
D. Student Sampling.....	6
E. Test Administration.....	6
F. Procedures.....	7
G. Technical Analysis.....	8
H. Sample Modifications.....	8
III. RESULTS - SUB-TESTS.....	8
A. List of Statistical Effects.....	8
B. Level I.....	10
1. Listening and Speaking.....	10
2. Reading.....	11
3. Writing.....	11
C. Level II.....	12
1. Listening.....	12
2. Reading.....	12
3. Speaking.....	12
4. Writing.....	13
D. Level III.....	13
1. Listening.....	13
2. Reading.....	13
3. Speaking.....	14
4. Writing.....	14

IV. RESULTS - TOTAL TEST SCORES.....	15
A. Introduction.....	15
B. Level I.....	15
C. Level II.....	15
D. Level III.....	16
E. Spring Pre-test and Spring Post-test.....	16
V. DISCUSSION.....	17
A. Introduction.....	17
B. Effects Attributable to Years in Program.....	17
C. Pre- to Post-test Differences.....	17
1. Spring-to-Spring District.....	17
2. Fall-to-Spring Sub-tests.....	18
3. Fall-to-Spring Total Test.....	18
VI. CONCLUSIONS.....	21
APPENDICES.....	28
REFERENCES.....	50

LIST OF TABLES AND FIGURES

Table 1 -	DFG: A and DFG: Not A Fall-Spring Testing...	5
Table 2 -	Mean Gains by Grade Level (in NCEs).....	20
Figure 1 -	Pre- to Post-test Gains in NCEs Total Test: By Years in Program.....	23
Figure 2 -	Pre- to Post-test Gains Total Test: By DFG Type.....	24
Figure 3 -	Pre- to Post-test Gains Total Test: By Grades.....	25
Figure 4 -	Total Test Gains by Grade Level (in NCEs)...	26
Figure 5 -	NCE Gains as a Function of Grades for Communication and Computation Basic Skills Preventive and Remedial.....	27

EXECUTIVE SUMMARY

Background

On January 8, 1975, the Governor signed the Bilingual Education Act to provide equal educational opportunity to students with limited English language skills. Bilingual education was to implement full-time instruction in: (1) all required courses, given in the native language of the enrolled students; (2) reading, writing, speaking and listening skills in both English and the native language of the students; (3) history and culture, both of the native land of the students' parents and of the United States.

Bilingual program participants receive both the needed instruction in English as a second language (ESL) and native language instruction in other areas. In this way, they avoid the educational delays which would normally occur if these students were assigned to classes conducted only in English. Under the bilingual strategy, students continue to learn their subjects in a language they can understand. As their English proficiency increases, instruction in other languages decreases. Thus, students are prepared for integration into English-speaking classrooms. Under normal conditions, such programs will be complete in three years.

When there are too few students for a bilingual program (20), students may only receive ESL instruction. Unfortunately, these students will continue to fall behind their classmates until they master the skills necessary to communicate with their English instructors, read English books, etc.

Purpose of Study

The current study was designed to determine whether students are acquiring English as a result of their program participation, whether they are catching up in English skills with English-speaking students over time,

and whether English acquisition depends on grade level, socioeconomic status, or other factors.

Procedures

To examine these concerns, the Language Assessment Battery (LAB, Riverside Publishing) was administered in the fall 1979 and again in the spring 1980, to a carefully selected sample of 731 students in bilingual programs. The major focus of testing was on Spanish-speaking students, to control for scoring variations which might arise from differences in languages, and also because this group represents the largest non-English linguistic population in the state.

In all, 731 bilingual students (grades one through twelve) were pre- and post-tested to assess English skills growth. Another 36 urban students, who had been pre-tested in the spring 1979, were post-tested in the spring 1980. Their scores are analyzed separately because they had been pre-tested six months before the other students in the study.

Student results were compared across grades, socioeconomic factors and numbers of years students had participated in bilingual programs. Analyses from pre-test to post-test also compared the progress of bilingual students with that of their English-speaking counterparts over the school year (in the case of the 36 urban students, this comparison was over the calendar year).

Results

In all grades tested, results consistently showed that third-year bilingual students significantly outscored first and second year students in English language skills. This supports the contention that bilingual education promotes English skills development.

It was also demonstrated that bilingual students had made significant gains relative to English-speaking children in eight of the twelve grades tested. For example, the average post-test score of second grade students who had started bilingual programs in kindergarten actually surpassed the average English-speaking students' theoretical norm (54.3 as compared to 50.0). This finding argues in favor of early bilingual education intervention.

Finally, it was determined that the amount of English growth was greatly influenced by grade level. Generally, elementary school students made much larger gains than high school students.

Recommendations

Based on these results, it is recommended that:

- Instruction in English language skills should begin as early as possible, and districts should uniformly screen the English proficiencies of incoming students to permit the earliest possible intervention;
- Bilingual education should include ESL intervention from the outset, if students are to realize their full potentials in English language skills. The data show that delayed ESL instruction limits student achievement.
- Districts should intensify English language experiences for secondary students who have had limited exposure to English and should be especially sensitive to the ESL, math and reading components of the secondary curricula.

I. INTRODUCTION

A. History

Since its adoption in 1976, the Bilingual Education Act has had significant impact on the education of those New Jersey students whose dominant language is not English. A major goal of the legislation is to ensure a "*thorough and efficient*" education for all students in all areas, regardless of language background. There now are approximately 120 bilingual programs in New Jersey. Districts have such programs for as many as seven language groups. These programs provide native language instruction in all academic areas while providing for the students' acquisition of English language skills.

B. Study Question: ~~Can Children Learn English While Being Instructed in~~ Their Native Language?

This study addresses the adequacy of these programs in promoting English proficiency. While this singular focus on language acquisition does not address all the goals of bilingual education, it does test whether children can be instructed in their dominant language at no cost to their acquisition of English. If they can, then bilingual education confers to them the advantage of not losing the instructional time they would normally lose while they learn English. However, if such instruction interferes with English acquisition, the students might profit more by a regular curriculum and English-as-a-Second-Language (ESL) instruction, even if they suffer an initial loss in learning due to lack of English mastery.

C. Research Designs

Before this study was undertaken, the answer to the question of whether children in bilingual education programs also are learning English

eluded researchers. Simple pre- and post-testing designs lacked sufficient statistical power to discern growth in English skills over the course of one year. Moreover, because bilingual populations are, by definition, different in language skills, control has always been lacking in these evaluation attempts. Finally, there is a dearth of valid instruments in the area, although New Jersey is continuing review of the psychometric adequacy of all available assessment instruments.

Fortunately, the transitional nature of programs in New Jersey, as prescribed by our legislation, administrative code and district practices, has enabled two alternative designs to be developed. The first, a modification of the Title I, Model A1 (Talmadge, 1976) permits a comparison of actual student English language growth with growth made by a comparable (age, grade) sample of monolingual English-speaking students over the same period of time. This is accomplished by assessing how the scores of bilingual students (relative to the test norms) change as a result of program intervention.

The second design is a between-subjects analysis, which is a simple, yet powerful research tool. This permits researchers to approximate three years of program effects by comparing the achievement of samples, randomly drawn from different program years. A recent analysis of students leaving bilingual programs gives us reason to believe that third-year students will significantly outscore second and first-year students.

D. Study Limitation

The Language Assessment Battery was developed by the New York City Board of Education. For the current study, it was administered in the fall and in the spring. Although national norms are available for this instrument for the spring scores, the fall and spring norms are based on a sample of children from New York City.

The current analysis relies on these latter norms. New Jersey's bilingual children are not actually compared to New York's children. Rather, one of the analyses tests whether they improve in their relative positions to this distribution of scores from pre-test to post-test.

II. METHODS AND PROCEDURES

A. Choice of Instruments

Two major criteria were employed to select a testing instrument:

1. The instrument must assess four areas of English language impact required by administrative code, i.e., reading, writing, speaking and listening; and
2. The instrument must have adequate psychometric properties, including demonstrated validity and reliability.

An exhaustive review of existing instruments recommended use of the Language Assessment Battery, (LAB, Riverside Publishing/New York City, 1976). Cut-off scores were not employed.

B. Determination of Sample Size

Each normed test was administered to a large sample of students. They were chosen because they represented the total population in terms of demography and ability levels. How these students scored determined the norms of the test. By comparison to the scoring distribution, any student's score may be converted to a percentile rank or a Normal Curve Equivalent, (NCE). These latter two scores, which range from one through 99, provide the relative position of the student in terms of the norming sample.

The distribution of LAB scores by the norming sample was used to

determine the number of students that would be tested. This number was computed by a formula that used two factors. The first, called "power", is the probability of finding true differences between groups of students. This was set at .80 in the present study. The second, called "alpha", is the probability that any differences found between groups could be attributed to chance fluctuations in scoring. This was limited to .05, which is an accepted statistical convention. These procedures, for instance, are what enable surveyors to generalize about political feelings, television preferences, etc., based on rather small segments of the total population. In this manner, the study could guarantee the most accurate results while limiting any inconvenience to the districts or students.

The minimum number of students required to meet these standards was 243 per year in program. Since third, second and first year students were tested, the total sample size required was 729. Additional students were pre-tested to account for high attrition rates and to permit teachers to test whole classes rather than to single out individual children.

C. District Sampling

Districts volunteered to participate on invitation from the Commissioner of Education. They were selected based on their having:

1. The requisite number of bilingual students in all grades from one through 12;
2. Fully implemented bilingual education programs for at least three years; and
3. A representative (to all current in New Jersey) instructional method.

The plan was to sample two districts each from the southern, central, northeastern, and northwestern regions. One would be of District Factor Group A

(DFG: A) and one of another District Factor Group (DFG: Not A).

District Factor Group (DFG) is a means of typifying school districts based on the socioeconomic status of average citizens within the districts. There are 11 DFGs, A to J and Z. "A" represents the lowest income status, while "J" is the highest. "Z" is reserved for those districts for which there is not enough information to make an adequate DFG assignment. Districts having bilingual programs are about evenly divided between Group A and other groups.

It was necessary to sample two Central Region districts of DFG: Not A to provide the requisite number of children at all grade levels because of the small size of programs in this region. Thus, nine districts participated. They tested their bilingual students in the fall of 1979 and again in the spring of 1980. A tenth district submitted data from a spring-to-spring test schedule, and that district's results were analyzed separately.

The participating districts were as follows:

TABLE 1

REGION	DFG: A	DFG: NOT A
NORTHEASTERN	EAST ORANGE	PLAINFIELD
NORTHWESTERN	PATERSON	DOVER
CENTRAL	NEW BRUNSWICK	CARTERET LAKEWOOD
SOUTH	ATLANTIC CITY	VINELAND

SPRING-TO-SPRING DISTRICT: UNION CITY (DFG: A)

D. Student Sampling

Districts were advised to sample whole classes whenever possible. These were randomly based on the school and teacher. The LAB is divided into three levels: Level I (Grades 1-2), Level II (Grades 3-6), and Level III (Grades 7-12).

Districts were asked to sample Spanish-speaking children only (to reduce possible variation in scores that could result from sampling different language groups). However, East Orange sampled many Haitian students and other districts included small numbers of Italian, Chinese, Arabic and Russian students. The effects of including these other children in the sample were proven to be statistically negligible before their data were aggregated in the total sample.

E. Test Administration

Two workshops on administration of the instrument were given in the fall and two more in the spring at the Educational Improvement Centers. District bilingual coordinators were invited to attend, along with an ESL teacher who would administer the instrument and, if available, a teacher who had administered the LAB before. County office representatives also attended.

The meetings were conducted by the Office of Evaluation and the Bureau of Bilingual Education. Sampling criteria, statistical design and purpose and projected outcomes were explained. A staff member from the Office of Policy Analysis demonstrated how to administer the LAB, provided the history of its development and pointed out its most valid uses. The staff member had been one of the original field administrators of the instrument in New York City.

Three supplemental meetings were held in Trenton at the Office of Evaluation to answer any further questions and to train districts unable to attend the original meetings.

To further ensure the validity of the chosen instrument for the current purpose:

1. The literature regarding the development of the instrument (the New York City Aspira Decree, etc.) was thoroughly reviewed.
2. The plan was reviewed with a number of other psychometric authorities at Riverside Publications, Educational Testing Service and the New York City Bureau of Evaluation.
3. The proposed use of the instrument was discussed with its authors.

F. Procedures

District ESL teachers administered the English portions of the LAB to their samples in October 1979 and April 1980. All materials were provided by the Office of Evaluation and collected after each administration.

New York's norms were used to convert total test scores to Normal Curve Equivalents (NCEs). Use of NCE scores enables statistical analyses that project how students would fare relative to an English-speaking population if they had not been in the bilingual education program. In this manner, growth was assessed relative to the English language growth of English-speaking children during the same time period. As mentioned earlier, NCEs reflect the achievement of students relative to the norming sample. They are often employed for analyses of this type because they have properties which enable them to be mathematically manipulated. These properties are not shared by percentiles or grade equivalents.

Subsections of the LAB are not normed. Therefore, analyses of subsections were conducted in raw scores.

G. Technical Analysis

The analysis used a general linear model: Years in Program (x); Grade (x); District Factor Grouping (x); Language for every level of the test. This analysis is basically a regression in which the effects of the above factors on scoring are compared to the effects of all other factors on scoring. The numbers reported in parentheses throughout the results section are group means.

There were two types of post hoc analyses. Scheffe contrasts and Duncan multiple range tests. These techniques are employed to pinpoint the exact nature of effects identified by the general linear models. For example, the general linear model may determine that grade level is a significant determinant of scoring. The post hoc analysis would show where these grade level differences occurred.

H. Sample Modifications

There were two small changes in the sample from the hand-edited count to the computer-edited count. Two additional students were added in the latter. The effect of including these two was negligible. Finally, one student's writing score on Level III was edited by the computer, making the *n* for the writing sub-test one less than the *n* for the other sub-tests.

III. RESULTS - SUB-TESTS

A. List of Statistical Effects

In the discussions that follow, group means will be given in parentheses and results will be analyzed in terms of statistical effects.

There are two types of effects, between subjects and within subjects. Between subjects refers to those variables for which every child can have only one value. For example, group membership is a between subjects variable. Children may either belong to a group of second graders or a group of third graders. They cannot belong to both of these groups. The second type of effect, within subjects, refers to those variables for which the child may have more than one value. For example, children may have two test scores each, one for pre-test and one for post-test.

The list of between subject and within subject variables used in the current analyses follows. An explanation is provided for each of these effects. Interactions between effects is described as one variable "by" another variable.

For example, the interaction of grade level with gains from pre-test to post-test would be described as Pre-Post by Grade.

BETWEEN SUBJECTS

Language	The language spoken by the child.
Years	The number of years (three, two or one) the child has been enrolled in the bilingual program.
Grade	The grade of the child.
DFG	District Factor Grouping: A socioeconomic variable. "A" is low and "Not A" is higher.

WITHIN SUBJECTS

Pre-Post	Gains made by the child from pre-test to post-test.
Error	All the variance in either between or within subjects analyses not accounted for by the above variables or their interactions.

B. Level I

1. Listening and Speaking (20 Items, Appendix 1)

There were significant differences attributed to District Factor Grouping, Years in Program and the Interaction of Grade and Years in program. Students from the districts in DFG: Not A (higher socio-economic status) scored a mean of 17.0, while those from DFG: A scored 16.0. The differences between pre- and post-testing were comparable for the two groups.

Post hoc tests demonstrated that the significant differences for Years in Program were between third- (18.8) and first-year groups (14.9), with the second year group (16.7) falling between the two. Post hoc analyses also revealed that second-graders who had been in the program for three years (18.8) significantly outscored first-year second graders (13.5). However, second-year first grade students (16.3) did not significantly outscore their first-year classmates (15.3).

There was a large, reliable difference in scoring from pre-test (15.7) to post-test (17.4). These differences were true of both DFG: A (means = 15.1 and 17.0 for pre- and post-tests, respectively) and DFG: Not A (means = 16.2 and 17.7, respectively). Significant pre- to post-test gains were made by first-year (6.7 to 8.1) and second-year students (8.0 to 8.7). Third year students also gained (9.2 to 9.6), but not significantly.

Finally, although there was a significant interaction involving Pre-Post by Years by Grade, the reliable differences involved groups that were not comparable (first-year second graders pre-test, 11.7 was significantly lower than third-year second graders' post-test, 19.2).

2. Reading (10 Items, Appendix 2)

DFG: Not A (mean = 5.8) scored significantly higher than DFG:A (mean = 4.8) and the second graders (7.1) outscored the first (3.7). Post hoc comparisons showed that Spanish-speaking children in both the first and second years of program participation (means = 4.2 and 5.2, respectively) scored significantly lower than those in their third program year (7.9). Furthermore, there was a tendency for third year students to outscore first-year students in both DFG: A (means = 3.8 for first- and third-year students), and in DFG: Not A (means = 4.8 for first and 6.8 for third-year students). Third-year students also outscored second year students in both types of districts (second-year means were 4.8 for DFG: A and 5.4 for DFG: Not A).

There were significant overall mean differences between pre-test (4.4) and post-test (6.2) scores. Furthermore, the Spanish-speaking children as a group (n = 242) made significant gains from a 4.5 pre-test mean to a 6.3 post-test mean. There also were significant rises both in DFG: A (pre-test mean = 3.8, post-test mean = 5.8) and DFG: Not A (pre-test mean = 5.0, post-test mean = 6.7). The gains were significant in both first (from 2.9 to 4.6) and second (from 6.2 to 8.1) grade. The significant Pre-Post by DFG by Years in program interaction also involved comparisons of no interest to the study (pre-test of first-year students in DFG: A, 2.4 compared to post-tests of third-year students in DFG: Not A, 9.6).

3. Writing (10 Items, Appendix 3)

Post hoc comparisons determined that second graders (7.6) scored significantly higher than first graders (4.5). Third-year students

also demonstrated significantly greater writing skills (8.2) than either first (5.0) or second-year students (5.7). The interaction of Grade by Years was less meaningful. First graders in their second program year (4.4) scored significantly lower than their second grade counterparts (7.6). Although there was a small pool of items, there was a significant gain in scores from pre-test (5.1) to post-test (6.8). A difference this large (33%) is most remarkable over the short time period.

C. Level II

1. Listening (30 Items, Appendix 5)

The grades scored in order as predicted: 21.8 (sixth), 21.5 (fifth), 21.0 (fourth), and 18.5 (third). Third-year students (21.7) significantly outscored first year students (18.5). Second-year students (20.7) fell between the two. There was also a significant difference between the pre-test mean (19.4) and post-test mean (21.9).

2. Reading (28 Items, Appendix 6)

Overall, the higher grades outscored the lower in reading; the sixth (16.3) and fifth (15.7) were significantly higher than the fourth (11.8). In turn, the fourth significantly outscored the third (8.9). Across grades, there was a significant difference between the pre-test mean (11.4) and post-test mean (14.7).

3. Speaking (14 Items, Appendix 7)

Language groups scored significantly different. The Haitian children (10.3) outscored the Spanish-speaking children (7.4). However, the number of children who did not speak Spanish (12) was much too small to permit meaningful conclusions. Once again, second (7.8) and third-year (8.7) students scored significantly higher than first-year students (5.1). Across years and grades, there was a significant difference between the

pre-test (6.1) and post-test means (8.9). The gain represents 46 percent of the pre-test score. The differences for both DFG: A (pre-test = 5.0, post-test = 8.8) and DFG: Not A (pre-test = 7.2, post-test = 9.0) were significant.

4. Writing (20 Items, Appendix 8)

True to the pattern of reading, sixth (13.9) and first (13.7) grade students scored significantly higher than fourth graders (10.6). Fourth graders also outscored third graders (8.1). An interaction of Grade by Years showed that third-year students in third grade (8.7) scored significantly lower than third-year sixth graders (15.5). Across grade levels, there was once again a significant difference between the pre-test mean (10.3) and post-test mean (12.7).

D. Level III

1. Listening (30 Items, Appendix 10)

The only significant difference between subjects effect for this variable was Language. The sample involved such small numbers of non-Spanish speaking children as to make the comparisons meaningless. Within subjects, there was again, a significant difference between the pre-test (16.4) and post-test means (17.6). There were also differences between the gains made by various language groups. While most groups gained one or two points, those who scored low on the pre-test made large improvements on the post-test.

2. Reading (28 Items, Appendix 11)

DFG: Not A (14.8) scored significantly higher than DFG: A (14.3). Also, the seventh graders (12.0) scored lower than eighth (14.9), ninth (14.9), tenth (16.5), eleventh (16.0) and twelfth (14.8) graders. Finally, students

in their third (15.7) and second-years (15.3) scored significantly higher than those in their first-year (12.7).

There was a significant gain across all students at this level, from the pre-test mean (13.3) to the post-test mean (15.6). There was also a significant gain for Spanish-speaking students as a group (pre-test = 13.4, post-test = 15.5). Significant gains were made by both the students of DFG: A (pre-test = 13.3, post-test = 15.3) and DFG: Not A (pre-test = 13.4, post-test = 16.2).

3. Speaking (14 Items, Appendix 12)

Although there once again was a significant effect for Language, the number of students speaking languages other than Spanish (12 again) was too small to permit generalizations. The effect for Years was very large. Third-year students (10.3) scored significantly higher than second-year students (9.0) or first-year students (6.8). Second-year students also significantly outscored first-year students. The DFG by Grade interaction involved combinations of means that were not of interest to the study.

4. Writing (20 Items, Appendix 13)

Once again, there were significant language effects and differences by grade levels. DFG: Not A (12.0) outscored DFG: A (11.2) and eleventh (13.8) and tenth graders (13.1) outscored eighth (10.5) and seventh graders (9.5). Twelfth graders (12.8) also scored significantly higher than seventh graders, while ninth graders (12.2) scored between the statistically differentiated groups. Third (12.3) and second-year (12.0) students also reliably outscored first-year students (10.2). The total sample also made a significant gain from pre-test (10.8) to post-test (12.1).

IV. RESULTS - TOTAL TEST SCORES

A. Introduction

Unlike the sub-tests, the total test scores were analyzed in Normal Curve Equivalents. Essentially, this allows the scores to reflect the student's performance relative to an English speaking population. Gains made are not just in absolute terms, but in relative terms. Thus, positive pre- to post-test rises mean that the student has gained ground on his/her English speaking peers by acquiring skills at an accelerated rate.

B. Level I (Appendix 4)

Third-year students achieved a mean (45.4) significantly higher than either second (33.6) or first-year students (27.3). The sample taken from DFG: Not A (37.2) also outscored those students from districts of DFG: A (30.2). Also, second-grade students who were in their first year of program participation (15.3) scored significantly lower than either their second (36.5) or third-year classmates (45.4). Overall, second graders rose significantly from a pre-test of 29.2 to a post-test of 43.4. A significant Language by pre-post effect revealed that a small number of Chinese-speaking children gained from a pre-test of 5.3 to a post-test of 29.0.

C. Level II (Appendix 9)

The third-year students (23.3) and second-year students (20.7) significantly outscored first-year students (13.4). Furthermore, students made significant progress relative to the English-speaking sample from pre-test (15.8) to post-test (24.0). Finally, there also was a significant interaction of pre-post by Grade by Years in program. However, post hoc comparisons showed that the differences were not of interest to the study questions.

D. Level III (Appendix 14)

Third-year students (17.7) scored significantly higher than first-year students (16.4). Second-year students (16.9) fell between the two, significantly different from neither.

E. Spring Pre-test and Spring Post-test

Union City submitted LAB data from a spring 1979 pre-test and spring 1980 post-test. On all three levels of the test, there was a significant effect for Years in Program:

1. On Level I, second- (23.1) and third-year (20.6) students scored significantly higher than first-year students (3.8).
2. On Level II, the same pattern of significance held true, while the order changed (third year = 24.5, second year = 13.8, and first year = 1.9).
3. On Level III, third-year students (25.1) significantly outscored both second (16.4) and first-year (2.8) students. Second-year students also scored significantly higher than first-year students.
4. On all three levels of the test, the post-tests were significantly higher than pre-tests:

<u>LEVEL</u>	<u>PRE-TEST MEAN</u>	<u>POST-TEST MEAN</u>	<u>NUMBER OF STUDENTS</u>
1	8.2	20.8	22
2	11.9	14.9	22
3	12.1	18.0	42

V. DISCUSSION

A. Introduction

The discussion must be organized in terms of the two major hypotheses:

(1) that children who have participated longer will exhibit greater English language skills; and (2) children in bilingual education programs will gain in English abilities relative to English-speaking children.

B. Effects Attributable to Years in Program

The results of the data analyzes supported the hypothesis that bilingual education in New Jersey is having an impact in terms of helping students acquire English language skills and hence preparing them to transition into a regular classroom situation. From the data, one can see that third-year students outscore first and/or second-year students which gives evidence and support to the fact that successful transition from Spanish to English is taking place. When examining the total test data using normal curve equivalent (NCE) scores, one can also conclude that progress is being made with regard to a comparable (grade, age) sample of non-bilingual English-speaking students.

Hypothetically, the bilingual students' abilities in English will reach the English skill levels of the norming group. In fact, on Level I of the test, the mean NCE of both pre- and post-test of 45.4 for third-year students is very close to the theoretical English-speaking mean of 50.0; moreover, the post-test mean for this group (54.3) actually surpasses the theoretical English-speaking norm. Obviously, three years of early intervention has dramatic effects on English skills.

C. Pre-to-Post-test Differences

1. Spring-to-Spring District

In the spring-to-spring data, there were significant gains from pre-test to post-test at all levels of the test. There is no evidence to suggest that these differences can be attributed to

the increased time from pre-test to post-test. Obviously, the English-speaking students to whom these children were compared would also benefit from increased time. The more likely conclusion is that the ESL component of this bilingual program encourages a rapid development of English language skills. There may be other districts in the sample having comparable rates of English acquisition.

2. Fall-to-Spring Sub-tests

Raw scores were used to analyze the sub-tests. Growth then is in an absolute sense. However, with the total test scores gains are relative to gains made by English-speaking students over the same time period.

In every sub-test, except Level III Speaking, there was significant growth in raw scores from pre-test to post-test. The students are making reliable progress over the year in all areas. With older students, surface speaking skills lag somewhat behind deeper linguistic skills, an interesting finding in view of psycholinguistic research that finds young children acquiring knowledge of linguistic rules before linguistic production increases.

3. Fall-to-Spring Total Test

The interpretation of the total test results are more complex. Findings may be somewhat conservative because there was very large mobility in this group. This even exceeded the overtesting strategy, resulting in loss of some power to detect pre- to post-test differences. While there were gains from pre-test to post-tests in NCEs at all levels of the LAB, these gains were only significant in Level II and

were at the "tendency" level, not exceeding usual levels of significance in Levels I and III.

The nature of bilingual education programs, as evidenced by the Years in Program effects, may further explain failure to detect reliable pre- to post-test NCE gains in Levels I and III. Simply stated, one year of program intervention usually does not have the sustained impact on English skills as multiple years.

Also, since each level of the test encompasses a band of grades, the instrument may not be as sensitive for the lowest grade in each level as later grades. In fact, the lowest mean NCE gains (-0.2 in grade 1 and -0.5 in grade 7) occurred in the first grades of Levels I and III. Grade 3 (Level II) was the exception to this observation. In grades 1 and 7, the NCE gains are at the random level. An analysis was made of each grade separately to determine if pre-test to post-test gains by test level obscured the gains made at each grade. Table 2 presents the results.

Table 2
MEAN GAINS BY GRADE LEVEL (in NCEs)

Grade	N	Pretest Mean	Posttest Mean	Mean Gain	T-Test
1	139	31.6	31.4	-0.2	0.01
2	125	29.2	43.4	14.2	6.38***
3	69	16.7	26.8	10.1	7.92***
4	61	17.6	25.5	7.9	7.00***
5	62	14.2	22.7	8.4	6.40***
6	63	14.7	21.0	6.3	4.41***
7	54	14.4	13.9	-0.5	0.30
8	51	14.6	16.1	1.5	1.10
9	35	5.8	10.5	4.7	4.20***
10	41	15.7	17.7	2.0	1.52
11	15	9.1	14.1	5.0	3.30***
12	16	7.1	11.1	4.0	2.56*

There were positive NCE gains in all grades other than first and seventh. In fact, eight of the ten positive gains were significant.

VI. CONCLUSIONS

The results of this study indicate that New Jersey's bilingual education program, in which students are instructed in their native language until they are proficient in English skills, effectively promotes development of these English skills. Added time in the program is related to increased abilities to read, write, speak and understand English.

Moreover, in early grades bilingual students do catch up, and even at times, surpass the English speaking students who comprised the test's norms. The degree of improvement was strongly related to grade level. This finding supports conclusions made in earlier psycholinguistic research, that a critical age exists for language acquisition. The earliest possible intervention maximizes student growth.

These data can be analyzed in view of what is known about skills acquisition in other curriculum areas in New Jersey. During the past two years, basic skills acquisition in both communication and computation has been related to grade level for the Basic Skills Preventive and Remedial population (DeMauro, 1980). A similar relationship between grade level and improvement applies (compare Figure 4, English Improvement to Figure 5, B.S.P. & R. Improvement).

These data suggest that immediate instruction in English, math and reading will greatly benefit the student because younger students tend to make larger relative gains than older. If these students' limited English abilities interfere with their ability to understand the instruction, the students may be deficient in the areas of instruction. On the other hand, delays in English instruction may put an artificial, avoidable ceiling on English skills.

Fortunately, instruction in one area need not preclude instruction in the other. Indeed, instructional models exist that provide:

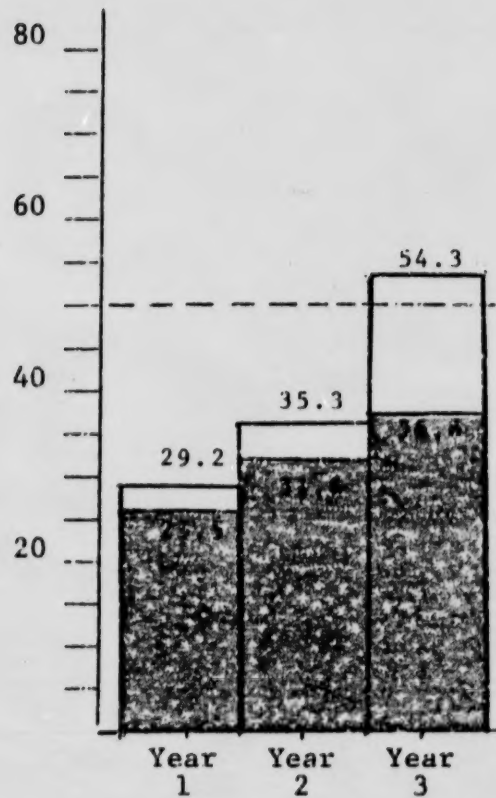
1. Early intervention in math and reading in a language that is comprehensible to the student;
2. Early intervention in English language skills; and
3. Even enrichment of native language skills.

Figure 1

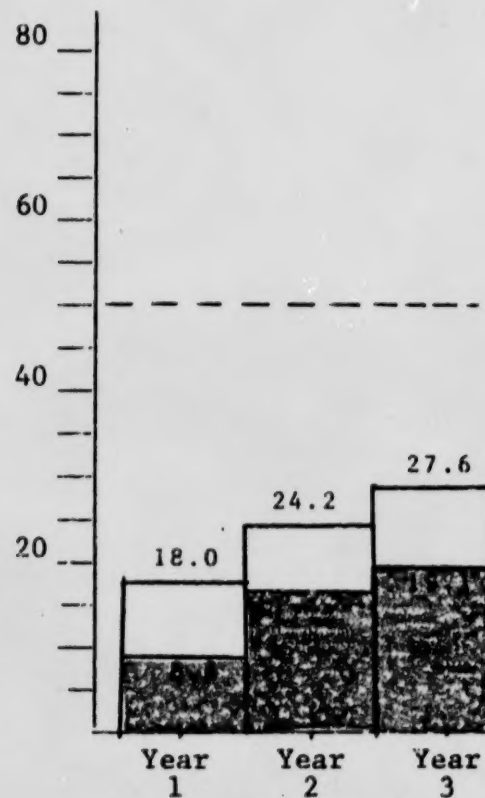
Pre to Posttest Gains in NCEs
Total Test: By Years in Program

Theoretical English-
Speaking Students' Mean -----

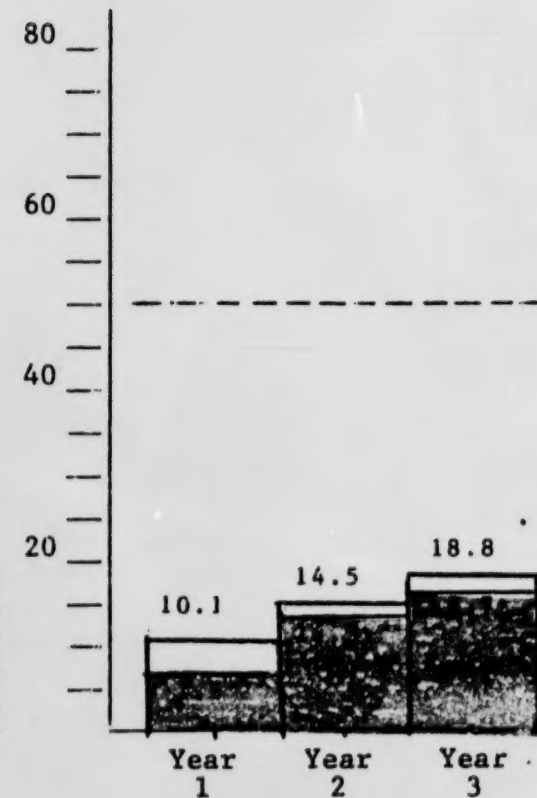
Pre ■
Post □



Level I



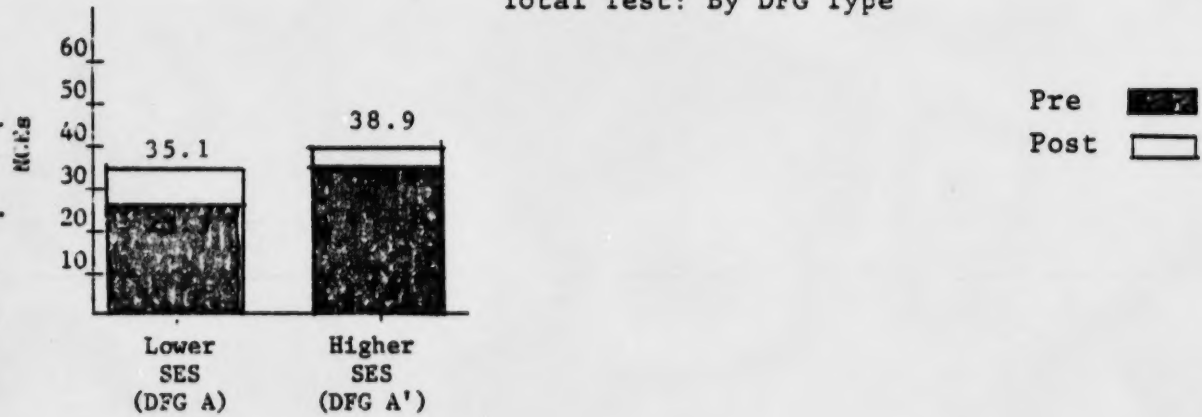
Level II



Level III

Figure 2

Pre to Posttest Gains
Total Test: By DFG Type



LEVEL I



LEVEL II



LEVEL III

Figure 3

Pre to Posttest Gains
Total Test: By Grades

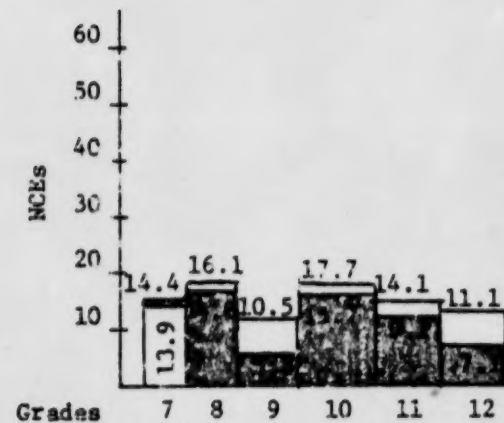
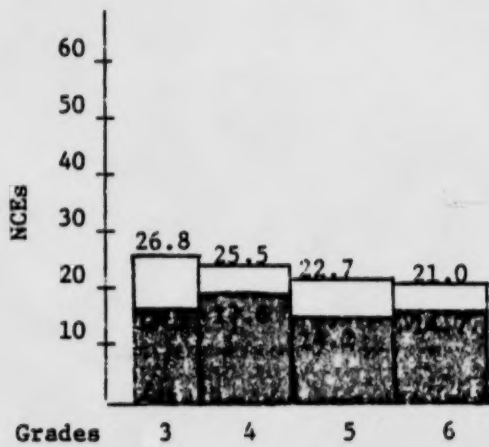
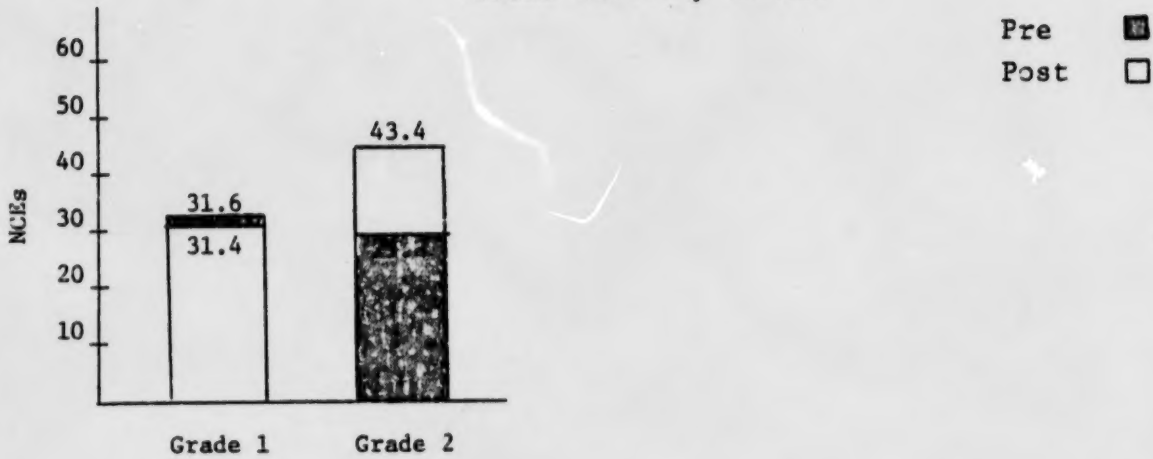


Figure 4

Total Test Gains by Grade Level
(in NCEs)

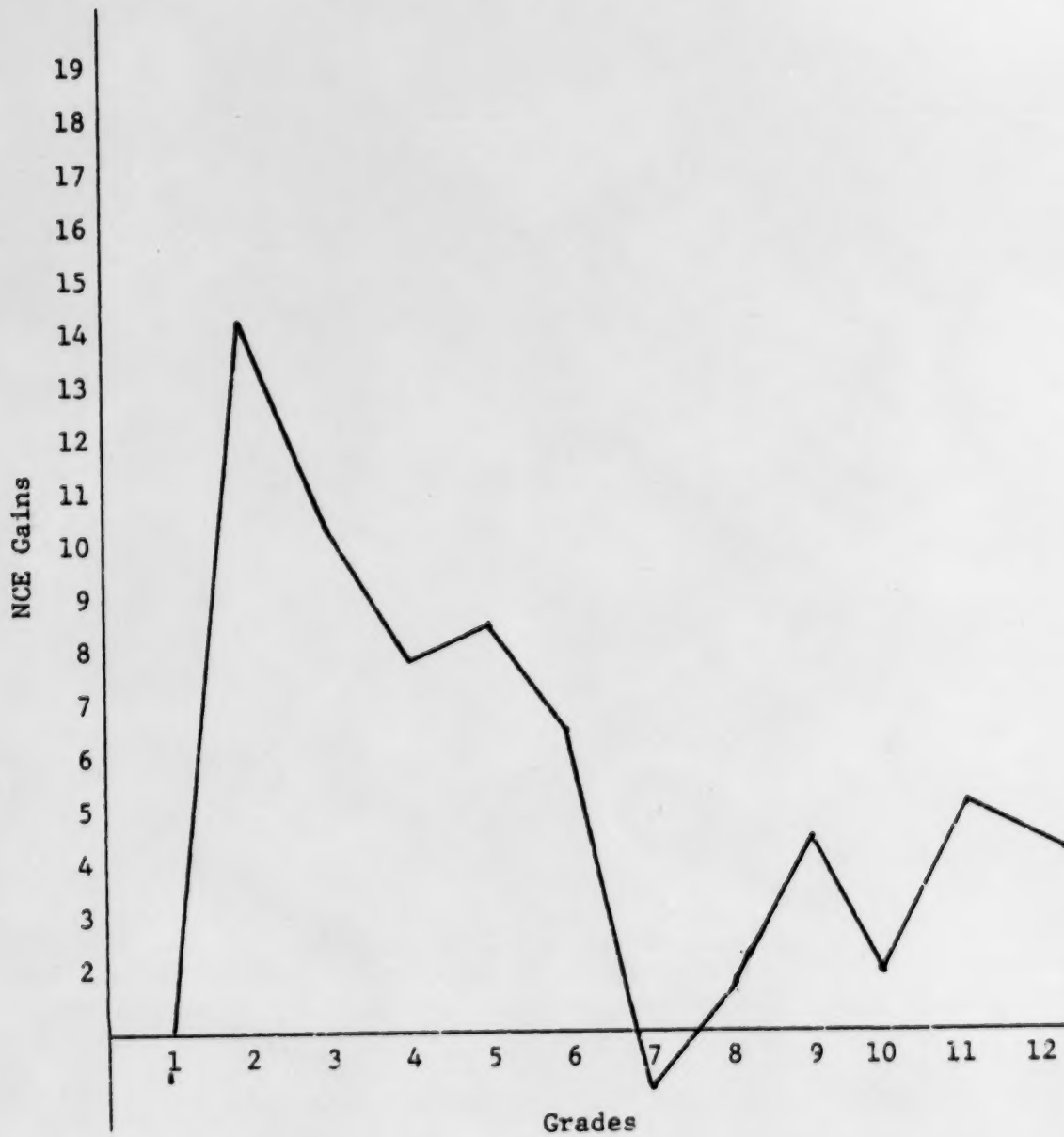
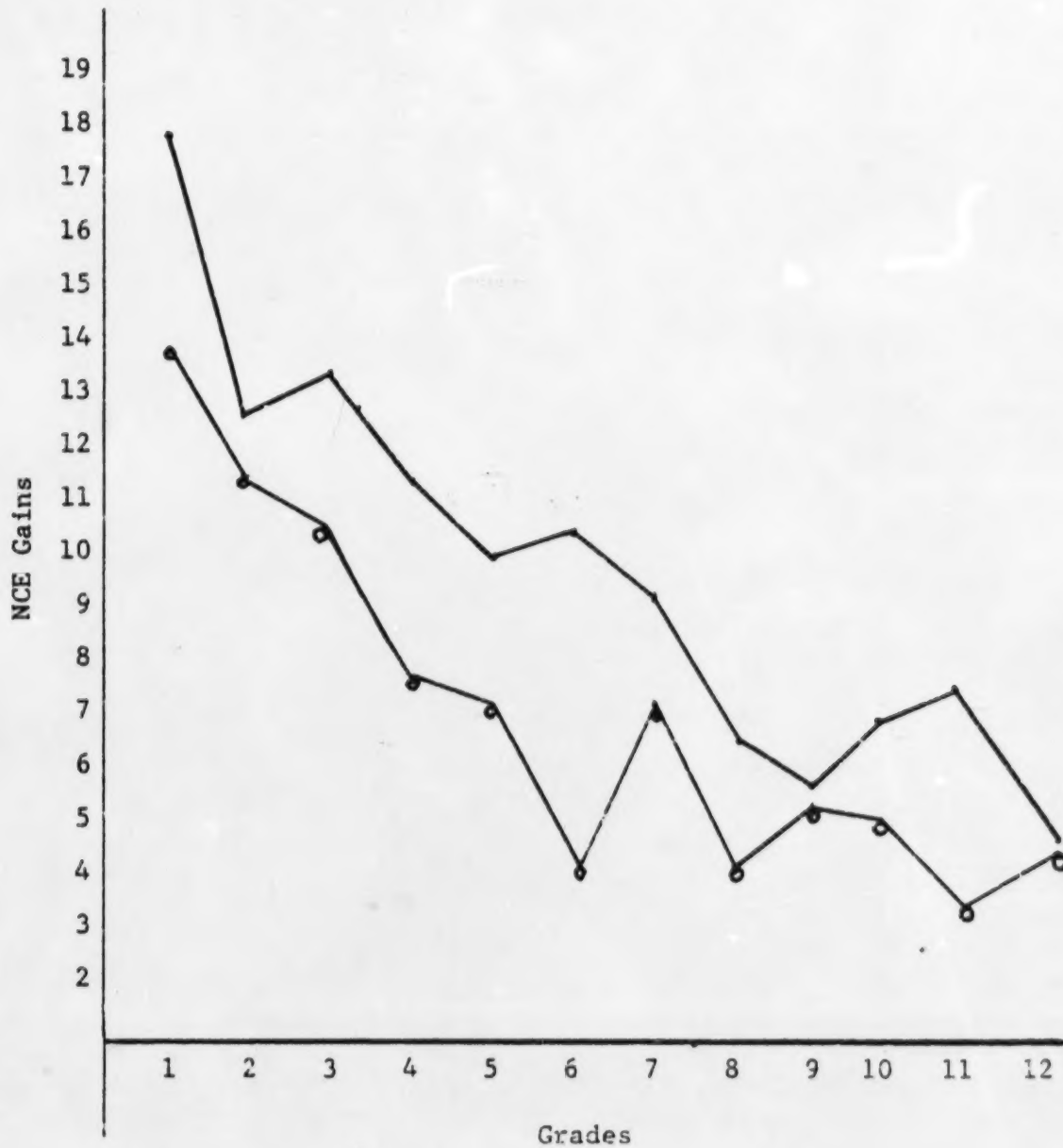


Figure 5

NCE Gains as a Function of Grades
for Communication and Computation
Basic Skills Preventive and Remedial



Note: small samples for grades 11 and 12

○ ——— ○ Comm.
● ——— ● Comp.

APPENDICES 1-17

Appendices 1-17 are statistical summary tables. The F-ratios provided in the last column tell the extent to which any given effect is significant. One star means an effect this great could occur by chance alone only 5 percent of the time. Two stars indicate that an effect this great could occur by chance alone only 1 percent of the time, and three stars indicate that such an effect could only occur by chance .1 percent of the time. The smaller the chance level, the larger the probability that the groups compared in the analysis (different grades, different DFGs, etc.) are truly different.

Appendix 1

Level I
Listening and Speaking
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	63.72	3	21.24	0.85
Years	402.12	3	134.04	5.39**
Grade	0.13	1	0.13	0.01
DFG	191.50	1	191.50	7.70**
Years*Grade	217.80	1	217.80	8.75**
DFG*Years	7.14	2	3.57	0.14
DFG*Grade	30.12	1	30.12	1.21
DFG*Years*Grade	26.55	1	26.55	1.07
Language*Years	41.45	3	13.82	0.56
Error	6145.33	247	24.88	
Total Between Ss.	<u>7125.86</u>	264		
Pre-Post	38.78	1	38.78	8.79**
Pre-Post*Lang.	7.55	3	2.52	0.57
Pre-Post*Yrs.	40.02	3	13.34	3.03*
Pre-Post*DFG	0.45	1	0.45	0.10
Pre-Post*Grade	5.92	1	5.92	1.34
Pre-Post*Yrs*Grade	17.27	1	17.27	3.91*
Pre-Post*DFG*Yrs.	10.09	2	5.04	1.14
Pre-Post*DFG*Grade	8.70	1	8.70	1.97
Pre-Post*DFG*Yrs*Grade	11.91	1	11.91	2.29
Grade Pre-Post*Lang*Yrs	3.56	3	1.19	0.29
Error	1075.64	244	4.41	
Total Within Ss.	<u>1219.89</u>	264		
Total	<u>8345.75</u>	527		

Appendix 2
Level I, Reading
ANOVA Summary Table

<u>Source</u>	<u>Sum Of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	23.79	3	7.93	0.87
Years	31.87	3	10.62	1.17
Grade	638.29	1	638.29	70.16***
DFG	185.56	1	185.56	20.39***
Years*Grade	42.34	1	42.34	4.65*
DFG*Years	1.75	2	0.88	0.10
DFG*Grade	14.96	1	14.96	1.64
DFG*Years*Grade	2.01	1	2.01	0.22
Language*Years	79.99	3	26.66	2.93*
Error	2247.15	247	9.10	
Total Between Ss.	3267.71	263		
Pre-Post	44.62	1	44.62	10.19***
Pre-Post*Lang.	55.69	3	18.56	4.24**
Pre-Post*Yrs.	11.06	3	3.69	0.84
Pre-Post*DFG	1.16	1	1.16	0.26
Pre-Post*Grade	3.85	1	3.85	0.88
Pre-Post*Yrs*Grade	0.97	1	0.97	0.22
Pre-Post*DFG*Yrs.	3.74	2	1.87	0.43
Pre-Post*DFG*Grade	11.67	1	11.67	2.66
Pre-Post*DFG*Yrs Grade	26.00	1	26.00	5.94*
Pre-Post*Lang*Yrs	4.68	3	1.56	0.36
Error	1081.96	247	4.38	
Total Within Ss.	1245.40	264		
Total	4513.11	527		

Appendix 3
Level I, Writing
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	9.79	3	3.26	0.52
Years	61.11	3	20.37	3.26*
Grade	505.43	1	505.43	80.86***
DFG	7.85	1	7.85	1.26
Years*Grade	52.08	1	52.08	8.33**
DFG*Years	27.73	1	13.87	2.22
DFG*Grade	13.23	2	13.23	2.12
DFG*Years*Grade	19.16	1	19.16	3.07
Language*Years	40.74	3	13.58	2.17
Error	1543.87	247	6.25	
Total Between Ss.	2280.99	263		
Pre-Post	39.38	1	39.38	17.51*
Pre-Post*Lang.	14.24	3	4.75	2.11
Pre-Post*Yrs.	0.51	3	0.17	0.08
Pre-Post*DFG	3.79	1	3.79	1.69
Pre-Post*Grade	6.48	1	6.48	2.88
Pre-Post*Yrs*Grade	1.27	1	1.27	0.56
Pre-Post*DFG*Yrs.	0.64	2	0.32	0.14
Pre-Post*DFG*Grade	2.55	1	2.55	1.13
Pre-Post*DFG*Yrs.*Grade	3.30	1	3.30	1.47
Pre-Post*Lang*Yrs	5.33	3	1.78	0.79
Error	555.50	247	2.25	
Total Within Ss.	632.99	264		
Total	2913.98	527		

Appendix 4
Level I, Total
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	473.97	3	157.99	0.15
Years	13547.34	3	4515.78	4.43**
Grade	687.42	1	687.42	0.67
DFG	8588.64	1	8588.64	8.42**
Years*Grade	11777.97	1	11777.97	11.55***
DFG*Years	234.34	2	117.17	0.11
DFG*Grade	862.06	1	862.06	0.85
DFG*Years*Grade	759.36	1	759.36	0.74
Language*Years	6805.47	3	2268.49	2.22
Error	251909.07	247	1019.87	
Total Between Ss.	<u>295645.64</u>	263		
Pre-Post	886.03	1	886.03	3.67
Pre-Post*Lang.	3535.06	3	1178.35	4.88**
Pre-Post*Yrs.	85.51	3	28.54	0.12
Pre-Post*DFG	768.82	1	768.82	3.18
Pre-Post*Grade	2832.72	1	2832.72	11.72***
Pre-Post*Yrs*Grade	26.19	2	13.20	0.05
Pre-Post*DFG*Yrs.	456.33	2	228.17	0.94
Pre-Post*DFG*Grade	776.89	2	388.45	1.61
Pre-Post*DFG*Yrs.*Grade	1007.14	1	1007.14	4.17*
Pre-Post*Lang*Yrs	751.62	3	250.54	1.04
Error	59677.55	247	241.61	
Total Within Ss.	<u>70803.86</u>	264		
Total	<u>366449.05</u>	527		

Appendix 5
Level II, Listening
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	111.35	2	55.68	1.20
Years	378.97	3	126.32	2.73
Grade	916.40	3	305.47	6.66***
DFG	11.68	1	11.68	0.25
Years*Grade	182.07	6	30.35	0.66
DFG*Years	24.45	2	12.22	0.26
DFG*Grade	82.85	3	27.62	0.60
DFG*Years*Grade	350.46	6	58.41	1.26
Language*Years	47.61	2	23.81	0.52
Error	10442.89	226	46.21	
Total Between Ss.	<u>12548.73</u>	254		
Pre-Post	42.74	1	42.74	4.07*
Pre-Post*Lang	8.10	2	4.05	0.39
Pre-Post*Yrs.	71.33	3	23.78	2.27
Pre-Post*DFG	7.67	1	7.67	0.73
Pre-Post*Grade	42.44	3	14.15	1.35
Pre-Post*Yrs*Grade	66.54	6	11.09	1.06
Pre-Post*DFG*Yrs.	20.91	2	10.46	1.00
Pre-Post*DFG*Grade	73.26	3	24.42	2.33
Pre-Post*DFG*Yrs*Grade	42.73	6	7.12	0.68
Pre-Post*Lang*Yrs*	9.15	2	4.58	0.44
Error	2371.11	226	10.49	
Total Within Ss.	<u>2755.68</u>	255		
Total	<u>15304.41</u>	509		

Appendix 6
Level II, Reading
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	169.64	2	84.82	1.43
Years	230.62	3	76.87	1.30
Grade	3558.72	3	1186.24	20.00**
DFG	33.81	1	33.81	0.57
Years*Grade	424.92	6	70.82	1.19
DFG*Years	25.74	2	12.87	0.22
DFG*Grade	165.59	3	55.20	0.93
DFG*Years*Grade	123.71	6	20.62	0.35
Language*Years	44.41	2	22.21	0.37
Error	13404.95	226	59.31	
Total Between Ss.	<u>18182.11</u>	254		
Pre-Post	188.38	1	188.38	20.21***
Pre-Post*Lang.	33.05	2	16.53	1.77
Pre-Post*Yrs.	17.50	3	5.83	0.63
Pre-Post*DFG	29.61	1	29.61	3.18
Pre-Post*Grade	7.82	3	2.61	0.28
Pre-Post*Yrs.*Grade	132.92	6	22.15	2.38
Pre-Post*DFG*Yrs.	42.36	2	21.43	2.30
Pre-Post*DFG*Grade	27.59	3	9.20	0.99
Pre-Post*DFG*Yrs*Grade	86.76	6	14.46	1.55
Pre-Post*Lang*Yrs	9.31	2	4.66	0.50
Error	2106.63	226	9.32	
Total Within Ss.	<u>2682.43</u>	255		
Total	<u>20864.54</u>	509		

Appendix 7
Level II, Speaking
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Squares</u>	<u>F-Ratio</u>
Language	223.09	2	111.55	5.61**
Years	215.99	3	72.00	3.62*
Grade	130.03	3	43.34	2.18
DFG	36.99	1	36.99	1.86
Years*Grade	81.47	6	13.58	0.68
DFG*Years	40.09	2	20.04	1.01
DFG*Grade	39.65	3	13.22	0.66
DFG*Years*Grade	118.68	6	19.78	0.99
Language*Years	3.10	2	1.55	0.08
Error	4496.04	226	19.89	
Total Between Ss.	<u>5385.13</u>	254		
Pre-Post	24.70	1	24.70	4.64*
Pre-Post*Lang	5.87	2	2.94	0.55
Pre-Post*Yrs.	41.45	3	13.82	2.60*
Pre-Post*DFG	121.40	1	121.40	22.82
Pre-Post*Grade	23.34	3	7.78	1.46
Pre-Post*Yrs*Grade	40.02	6	6.67	1.25
Pre-Post*DFG*Yrs.	7.27	2	3.64	0.68
Pre-Post*DFG*Grade	3.84	3	1.28	0.24
Pre-Post*DFG*Yrs.*Grade	21.98	6	3.66	0.69
Pre-Post*Lang*Yrs	5.74	2	2.87	0.54
Error	1202.76	226	5.32	
Total Within Ss.	<u>1498.37</u>	255		
Total	<u>6883.50</u>	509		

Appendix 8

Level II, Writing
ANOVA Summary Table

<u>Source</u>	<u>Sum of Square</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	160.74	2	80.37	1.99
Years	127.70	3	42.57	1.05
Grade	1983.47	3	631.16	16.34***
DFG	78.62	1	78.62	1.94
Years*Grade	550.85	6	91.81	0.74
DFG*Years	42.74	2	21.37	0.53
DFG*Grade	31.72	3	10.54	0.26
DFG*Years*Grade	180.18	6	30.03	0.74
Language	19.28	2	9.64	0.24
Error	9143.58	226	40.46	
Total Between Ss.	<u>12318.88</u>	<u>254</u>		
Pre-Post	59.37	1	59.37	7.01**
Pre-Post*Lang.	5.83	2	2.92	0.34
Pre-Post*Yrs.	11.67	3	3.89	0.46
Pre-Post*DFG	0.02	1	0.02	0.00
Pre-Post*Grade	31.67	3	10.56	1.25
Pre-Post*Yrs*Grade	94.70	6	15.78	1.86
Pre-Post*DFG*Yrs.	2.35	2	1.18	0.14
Pre-Post*DFG*Grade	8.04	3	2.68	0.32
Pre-Post*DFG*Yrs*Grade	60.36	6	10.06	1.19
Pre-Post*Lang*Yrs	1.63	2	0.82	0.10
Error	1912.83	226	8.46	0.10
Total Within Ss.	<u>2188.47</u>	<u>255</u>		
Total	<u>14507.35</u>	<u>509</u>		

Appendix 9
Level II, Total Test
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	1095.90	2	547.95	1.39
Years	4197.00	3	1399.00	3.54 *
Grade	1654.77	3	551.59	1.40
DFG	132.23	1	132.23	0.33
Years*Grade	1389.43	6	231.57	0.59
DFG*Years	458.43	2	229.22	0.58
DFG*Grade	1353.78	3	451.26	1.14
DFG*Years*Grade	1629.15	6	271.53	0.69
Language*Years	1134.60	2	567.30	1.44
Error	89208.97	226	394.73	
Total Between Ss.	102254.26	254		
Pre-Post	1400.48	1	1400.48	27.16***
Pre-Post*Lang	195.14	2	97.57	1.89
Pre-Post*Yrs.	254.36	3	84.79	1.64
Pre-Post*DFG	31.12	1	31.12	0.64
Pre-Post*Grade	303.96	3	101.32	1.96
Pre-Post*Yrs.*Grade	765.93	6	127.66	2.48*
Pre-Post*DFG*Yrs.	197.35	2	98.68	1.91
Pre-Post*DFG*Grade	46.38	3	15.46	0.30
Pre-Post*DFG*Yrs*Grade	316.34	6	52.72	1.02
Pre-Post*Lang*Yrs	277.36	2	138.68	2.69
Error	11654.70	226	51.57	
Total Within Ss.	15443.12	255		
Total	117697.35	509		

Appendix 10
Level III, Listening
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	531.11	6	88.52	3.18**
Years	102.90	2	51.45	1.85
Grade	72.39	5	14.48	0.52
DFG	0.48	1	0.48	0.02
Years*Grade	186.82	10	18.68	0.67
DFG*Years	123.13	2	61.57	2.21
DFG*Grade	212.24	5	42.45	1.52
DFG*Years*Grade	304.75	9	33.86	1.22
Language*Years	35.68	2	17.84	0.64
Error	4704.76	169	27.83	
Total Between Ss.	<u>6274.26</u>	211		
Pre-Post	85.79	1	85.79	10.48**
Pre-Post*Lang.	114.75	6	19.13	2.34*
Pre-Post*Yrs.	19.14	2	9.57	1.17
Pre-Post*DFG	0.36	1	0.36	0.04
Pre-Post*Grade	59.52	5	11.90	1.45
Pre-Post*Yrs*Grade	58.39	10	5.84	0.71
Pre-Post*DFG*Yrs	21.72	2	10.86	1.33
Pre-Post*DFG*Grade	34.66	5	6.93	0.84
Pre-Post*DFG*Yrs*Grade	42.64	9	4.74	0.58
Pre-Post*Lang*Yrs	15.62	2	7.81	0.95
Error	1833.88	169	8.19	
Total Within Ss.	<u>2136.47</u>	212		
Total	<u>8410.73</u>	423		

Appendix 11
Level III, Reading
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	484.59	6	80.77	1.85
Years	287.34	2	143.67	3.29*
Grade	1213.86	5	242.77	5.55***
DFG	224.01	1	224.01	5.12*
Years*Grade	311.16	10	31.11	0.71
DFG*Years	180.90	2	90.45	2.07
DFG*Grade	357.92	5	71.58	1.64
DFG*Years*Grade	438.03	9	48.67	1.11
Language*Years	114.39	2	57.20	1.31
Error	7388.09	169	43.72	
Total Between Ss.	<u>11000.29</u>	211		
Pre-Post	114.42	1	114.42	13.06***
Pre-Post*Lang.	137.89	6	22.98	2.62*
Pre-Post*Yrs.	37.77	2	18.88	2.16
Pre-Post*DFG	35.88	1	35.88	4.10*
Pre-Post*Grade	38.37	5	7.67	0.88
Pre-Post*Yrs*Grade	95.41	10	9.54	1.09
Pre-Post*DFG*Yrs.	65.87	2	32.93	3.76*
Pre-Post*DFG*Grade	59.58	5	11.92	1.36
Pre-Post*DFG&Yrs*Grade	83.81	9	9.31	1.06
Pre-Post*Lang*Yrs	9.44	2	4.72	0.54
Error	1480.43	169	8.76	
Total Within Ss.	<u>2158.87</u>	212		
Total	<u>13159.16</u>	423		

Appendix 12
Level III, Speaking
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	334.14	6	55.83	2.41*
Years	175.35	2	87.68	3.79*
Grade	216.01	5	43.20	1.87
DFG	6.06	1	6.06	0.26
Years*Grade	163.67	10	16.37	0.71
DFG&Years	50.48	2	25.24	1.09
DFG&Grade	266.15	5	53.23	2.30*
DFG*Years*Grade	177.14	9	19.68	0.85
Language*Years	23.09	2	11.55	0.50
Error	3911.36	169	23.14	
Total Between Ss.	<u>5323.45</u>	211		
Pre-Post	5.25	1	5.25	0.66
Pre-Post&Lang	18.69	6	3.12	0.39
Pre-Post*Yrs.	13.45	2	6.73	0.84
Pre-Post*DFG	18.19	1	18.19	2.28
Pre-Post*Grade	52.30	5	10.46	1.31
Pre-Post*Yrs*Grade	36.34	10	3.63	0.46
Pre-Post*DFG*Yrs.	6.43	2	3.21	0.40
Pre-Post*DFG*Grade	17.29	5	3.46	0.43
Pre-Post*DFG*Yrs*Grade	109.38	9	12.15	1.52
Pre-Post*Lang*Yrs	3.34	2	1.67	0.21
Error	1348.18	169	7.98	
Total Within Ss.	<u>1628.84</u>	212		
Total	<u>6952.29</u>	423		

Appendix 13
Level III, Writing
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Language	726.99	6	121.17	3.60**
Years	193.94	2	96.97	2.88
Grade	853.23	5	170.65	5.07***
DFG	217.59	1	217.59	6.46*
Years*Grade	263.39	10	26.34	0.78
DFG*Years	56.58	2	28.29	0.84
DFG*Grade	212.33	5	42.47	1.26
DFG*Years*Grade	345.65	9	38.41	1.14
Language*Years	88.32	2	44.16	1.31
Error	5658.71	168	33.68	
Total Between Ss.	8616.73	210		
Pre-Post	60.87	1	60.87	5.65*
Pre-Post*Lang.	130.86	6	21.81	2.02
Pre-Post*Yrs.	17.40	2	8.70	0.81
Post-Post*DFG	1.12	1	1.12	0.10
Pre-Post*Grade	35.24	5	7.05	0.65
Pre-Post*Yrs*Grade	55.29	10	5.53	0.51
Pre-Post*DFG*Yrs.	19.54	1	9.77	0.91
Pre-Post*DFG*Grade	50.94	5	10.19	0.95
Pre-Post*DFG*Yrs*Grade	69.82	9	7.76	0.72
Pre-Post*Lang*Yrs	24.72	2	12.36	1.15
Error	1820.96	169	10.77	
Total Within Ss.	2286.76	211		
Total	10903.49	421		

Appendix 14
Level III, Total Test
ANOVA Summary Table

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
Language	1742.37	6	290.39	1.03
Years	2392.34	2	1196.17	4.25*
Grade	1905.22	5	381.04	1.35
DFG	203.63	1	203.63	0.72
Years*Grade	1351.08	10	135.11	0.48
DFG*Years	899.18	2	449.59	1.60
DFG*Grade	1477.50	5	295.50	1.05
DFG*Years*Grade	2906.85	9	322.98	1.15
Language*Years	787.15	2	393.58	1.40
Error	47545.49	169	281.33	
Total Between Ss.	61210.80	211		
Pre-Post	55.75	1	55.75	1.21
Pre-Post*Lang.	204.71	6	34.12	0.74
Pre-Post*Yrs.	9.18	2	4.59	0.10
Pre-Post*DFG	22.70	1	22.70	0.49
Pre-Post*Grade	262.03	5	52.41	1.14
Pre-Post*Yrs*Grade	184.30	10	18.30	0.40
Pre-Post*DFG*Yrs.	24.53	2	12.26	0.27
Pre-Post*DFG*Grade	122.55	5	24.51	0.53
Pre-Post*DFG*Yrs*Grade	347.04	9	38.56	0.84
Pre-Post*Lang*Yrs	93.38	2	46.69	1.02
Error	7762.97	169	45.93	
Total Within Ss.	9089.14	212		
Total	70299.94	423		

Appendix 15
Union City, Level I
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Years	3542.04	2	1771.02	15.46***
Error	2175.96	19	114.52	
Total Between Ss.	5718.00	21		
Pre-Post	1857.62	1	1857.62	15.56***
Pre-Post*Yrs.	543.77	2	271.89	2.28
Error	2267.95	19	119.36	
Total Within Ss.	4669.34	22		
Total	10387.34	43		

Appendix 16
Union City, Level II
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Years	1434.82	2	717.41	3.19*
Grade	562.71	2	281.36	1.25
Years*Grade	75.54	2	37.77	0.17
Error	3371.11	15	224.74	
Total Between Ss.	<u>5444.18</u>	21		
Pre-Post	176.93	1	176.93	6.75*
Pre-Post*Yrs.	100.81	2	50.40	1.92
Pre-Post*Grade	192.38	2	96.19	3.67
Pre-Post*Yrs*Grade	13.04	2	6.52	0.25
Errors	392.97	15	261.98	
Total Within Ss.	<u>876.13</u>	22		
Total	<u>6320.31</u>	43		

Appendix 17
Union City, Level III
ANOVA Summary Table

<u>Source</u>	<u>Sum of Squares</u>	<u>Degrees of Freedom</u>	<u>Mean Square</u>	<u>F-Ratio</u>
Years	6038.69	2	3019.34	24.56**
Grade	1189.13	4	297.28	2.42
Years*Grade	1135.28	8	141.91	
Error	3319.38	27	122.94	
Total Between Ss.	<u>11682.48</u>	41		
Pre-Post	517.44	1	517.44	10.26**
Pre-Post*Yrs.	138.00	2	69.00	1.37
Pre-Post*Grade	363.04	4	90.76	1.80
Pre-Post*Yrs*Grade	149.69	8	18.71	0.37
Error	1361.70	27	50.43	
Total Within Ss.	<u>2529.87</u>	43		
Total	<u>14212.35</u>	83		

MIS 05C00/
Exp. Date 6/30/80

(FOR STATE USE ONLY)

--	--

--	--	--	--



(COLUMNS 6 - 11 FOR STATE USE ONLY)

-04-

DATE _____

55

Appendix 19

District Sampling for English
Proficiency Testing, Fall 1979

Districts were selected for assessment on a variety of uniformly applied criteria, based on data made available by the Bureau of Bilingual Education. Briefly, these criteria and the rationales are listed below:

<u>Criteria</u>	<u>Rationale</u>
1. Bilingual program at least three years old.	To enable representation of an adequate number of children in each of the three years of operation as outlined in the Administrative Code.
2. Adequate sample size in each grade level.	To ensure adequate power of grade level comparisons as projected by the standard error of the instrument.

Stratifying Criteria

3. Region	To ensure proper representation of the state's target population.
4. DFG	As a base for equating with the normative population of the instrument and to discern differential scoring.

Once a list of possible districts meeting these requirements was compiled, a number of other factors were considered, both to reduce noise in the procedures and also to ensure representativeness of the sample. Among these were curriculum variables.

Most bilingual programs in the state now are in districts in which there are a variety of language groups. Some of these groups have fewer children, and therefore receive ESL, rather than full bilingual programs. These programs are, for the most part, pullouts, while the bilingual programs have different educational environments.

The variety of programs in New Jersey is great and it would be difficult to assess any state limited English population across all types. The district sample was chosen for representativeness rather than control. It will be possible, however, to assess the differential effects of the sampled educational environment.

<u>Total New Jersey</u>	<u>Sampled Districts</u>	<u>Description</u>
.085	.100	Pullout only
.191	.100	Learning Center/ Self-contained class
.277	.400	Combination of above
.447	.400	Other/Multiple combinations

References

- Barr, A.J., Goodnight, J.H., Sall, J.P., & J.T. Helwig, A User's Guide to SAS 79. Raleigh, North Carolina: SAS Institute, Inc., 1976.
- Danoff, M.N., Evaluation of the Impact of ESEA Title VII Spanish/English Bilingual Education Program: Overview of Study and Findings. Washington, D.C.: Office of Planning, Budgeting, and Evaluation, Office of Education, 1978.
- DeMauro, G.E., Basic Skills Preventive and Remedial Programs, 1978-1980. Trenton, New Jersey: Office of Evaluation, N.J. State Department of Education, 1980.
- Myers, J.L., Fundamentals of Experimental Design (Second Edition). Boston: Allyn and Bacon, Inc., 1972.
- Tallmadge, G.K., Interpreting NCEs: ESEA Title I Evaluation and Reporting System, Technical Paper No. 2, Mountain View, California: RMC Corporation, 1976.
- Tallmadge, G.K. & C.T. Wood, User's Guide: ESEA Title I Evaluation and Reporting System. Mountain View, California: RMC Corporation, 1976.